

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
ATTORNEY DOCKET NO. 047912-0136

j1046 U.S. PRO
09/878270
06/12/01

Applicant: Kenichi MIKI
Title: DATA READ/WRITE CONTROLLING METHOD, DISK ARRAY
APPARATUS, AND RECORDING MEDIUM FOR RECORDING
DATA READ/WRITE CONTROLLING PROGRAM
Appl. No.: Unassigned
Filing Date: 06/12/2001
Examiner: Unassigned
Art Unit: Unassigned

CLAIM FOR CONVENTION PRIORITY

Commissioner for Patents
Washington, D.C. 20231

Sir:


The benefit of the filing date of the following prior foreign application filed in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of said original foreign application:

Japanese Patent Application No. 2000-177036 filed June 13, 2000.

Respectfully submitted,

June 12, 2001
Date



David A. Blumenthal
Attorney for Applicant
Registration No. 26,257

FOLEY & LARDNER
Washington Harbour
3000 K Street, N.W., Suite 500
Washington, D.C. 20007-5109
Telephone: (202) 672-5407
Facsimile: (202) 672-5399

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

miki
047912/0136

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

2000年 6月13日

出願番号
Application Number:

特願2000-177036

出願人
Applicant(s):

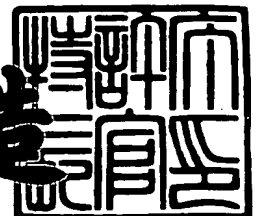
日本電気株式会社



2001年 2月23日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2001-3011533

【書類名】 特許願

【整理番号】 67000013

【提出日】 平成12年 6月13日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F

【発明の名称】 データ読み書き制御方法及びディスクアレイ装置並びに
データ読み書き制御用プログラムを記憶した記録媒体

【請求項の数】 9

【発明者】

 【住所又は居所】 東京都港区芝五丁目7番1号 日本電気株式会社内

 【氏名】 三木 健一

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100079164

 【弁理士】

 【氏名又は名称】 高橋 勇

 【電話番号】 03-3862-6520

【手数料の表示】

 【予納台帳番号】 013505

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9003064

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ読み書き制御方法及びディスクアレイ装置並びにデータ読み書き制御用プログラムを記憶した記録媒体

【特許請求の範囲】

【請求項 1】 ディスクに書き込まれる所定のデータを上位ホストから受け取るデータ受け取り工程と、当該受け取ったデータを所定の処理をするデータ処理工程と、当該処理したデータをディスクに書き込むデータ書き込み工程とを備え、

前記データ処理工程は、前記データ受け取り工程で受け取ったデータを複数に分割すると共にパリティデータを生成するデータ分割工程と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納工程と、前記各分割データとパリティデータとを前記キャッシュモジュールから取り出すと共にいずれかの分割データが破損している場合に当該分割データを前記パリティデータを用いて修復するデータ修復工程と、これら分割データを合成するデータ合成工程とを備えたことを特徴とするデータ読み書き制御方法。

【請求項 2】 上位ホストに送信する所定のデータをディスクから読み出すデータ読み出し工程と、当該読み出したデータを所定の処理をするデータ処理工程と、当該処理したデータを上位ホストに送信するデータ送信工程とを備え、

前記データ処理工程は、前記データ読み出し工程で読み出したデータを複数に分割すると共にパリティデータを生成するデータ分割工程と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納工程と、前記各分割データとパリティデータとを前記キャッシュモジュールから取り出すと共にいずれかの分割データが破損している場合に当該分割データを前記パリティデータを用いて修復するデータ修復工程と、これら分割データを合成するデータ合成工程とを備えたことを特徴とするデータ読み書き制御方法。

【請求項 3】 上位ホストからの指令によりディスクに対して所定のデータを書き込み及び読み出しすると共に、当該所定のデータを演算処理するアレイ制御部を備えたディスクアレイ装置において、

前記アレイ制御部は、前記所定のデータを少なくとも 2 つに分割すると共に当

該所定のデータのパリティデータを生成するデータ分割機能と、前記分割データ1つが破損している場合に前記パリティデータを用いて当該破損した分割データを修復する共に前記分割データを合成するデータ合成機能とを備えたことを特徴とするディスクアレイ装置。

【請求項4】 上位ホストからの指令によりディスクに対して所定のデータを書き込み及び読み出しすると共に、当該所定のデータを演算処理するアレイ制御部を備えたディスクアレイ装置において、

前記アレイ制御部は、前記所定のデータを少なくとも2つに分割すると共に当該所定のデータに基づいてパリティデータを生成するデータ分割部と、前記各分割データと前記パリティデータとをそれぞれ一時的に記憶する複数のキャッシュモジュールと、前記キャッシュモジュールの1つが故障した場合に当該故障のキャッシュモジュールに格納されていた分割データを残りの分割データと前記パリティデータとを用いて修復すると共に前記分割データを合成するデータ合成部とを備えたことを特徴とするディスクアレイ装置。

【請求項5】 前記各キャッシュモジュールの容量を、それぞれ等しく設定したことを特徴とする請求項4記載のディスクアレイ装置。

【請求項6】 前記各分割データ及びパリティデータの容量を、それぞれ等しく設定したことを特徴とする請求項4又は5記載のディスクアレイ装置。

【請求項7】 前記分割データ数と前記パリティデータ数との合計数を、前記キャッシュモジュール数に等しく設定したことを特徴する請求項4，5又は6記載のディスクアレイ装置。

【請求項8】 前記分割データ数を、前記キャッシュモジュール数より1つ少ない個数に設定したことを特徴とする請求項4，5，6又は7記載のディスクアレイ装置。

【請求項9】 ディスクに書き込まれる所定のデータを上位ホストから受け取った後、当該データを複数に分割すると共にパリティデータを生成するデータ分割処理と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納処理と、前記各分割データとパリティデータとを前記キャッシュモジュールから取り出すと共にいずれかの分割データに破損があ

る場合に当該破損している分割データを前記パリティデータとを用いて修復するデータ修復処理と、これら分割データを合成すると共に当該データをディスクに書き込むデータ合成処理とを備えたことを特徴とするデータ読み書き制御用プログラムを記憶した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ディスクアレイ装置に係り、特にデータに冗長性を持たせ、当該データをディスクに格納するディスクアレイ装置に関する。

【0002】

【従来の技術】

従来より、ディスクドライブに対するデータの書き込み及び読み出し速度の向上のため、キャッシュメモリを介して、上位ホストとディスクとのデータの転送を行っている。

【0003】

そして、従来のディスクアレイ装置では、さらにデータ保存の信頼性向上のため、特開平11-312058に見られるように、キャッシュ内に同一のデータをコピーし、どちらか一方のキャッシュが破損しても、他方のキャッシュ内のデータを転送するという方法が採られている。

【0004】

具体的には、キャッシュの冗長化の実現方法として、図3のような構成をとっている。この構成によるキャッシュの冗長化では、キャッシュモジュール109とキャッシュモジュール110とに、内部データバス112を介して同一のデータを互いにコピー（ミラー化）しておくことにより実現している。

【0005】

【発明が解決しようとする課題】

しかしながら、前述した方法では、同一のデータをミラー化するため、データの格納容量は実装キャッシュ容量の2分の1しか割り当てることができず、大容量のデータの転送ができないという問題が生じる。

【0006】

また、キャッシュモジュールのいずれか1モジュールが故障した場合において、引き続き残りのキャッシュモジュールでキャッシュ機能を使用することができず、ホストコンピュータに対するI/O性能の低下を招くという問題も生じる。

【0007】

【発明の目的】

本発明は、かかる従来例の有する不都合を改善し、特にキャッシュが故障した場合においてもデータを修復しデータ転送の信頼性を保つと共に実装キャッシュ容量のデータ格納容量の割合を増加してデータ転送の大容量化を図り、ホストコンピュータに対するI/O性能の低下を防ぐことができるデータ読み書き制御方法及びディスクアレイ装置並びにデータ読み書き制御用プログラムを記憶した記録媒体を提供することをその目的とする。

【0008】

【課題を解決するための手段】

上記目的を達成するため、請求項1記載の発明では、ディスクに書き込まれる所定のデータを上位ホストから受け取るデータ受け取り工程と、当該受け取ったデータを所定の処理をするデータ処理工程と、当該処理したデータをディスクに書き込むデータ書き込み工程とを備えている。

【0009】

そして、データ処理工程は、データ受け取り工程で受け取ったデータを複数に分割すると共にパリティデータを生成するデータ分割工程と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納工程と、各分割データとパリティデータとをキャッシュモジュールから取り出すと共にいずれかの分割データが破損している場合に当該分割データをパリティデータを用いて修復するデータ修復工程と、これら分割データを合成するデータ合成工程とを備える、という構成を採っている。

【0010】

このため、請求項1記載の発明では、上位ホストから受け取ったデータをまず複数に分割して、当該データのパリティデータを生成する。これら各分割データ

とパリティデータとをそれぞれ個別にキャッシュモジュールに格納する。そして、キャッシュモジュールから前記各データを取り出し、いずれかの分割データが破損している場合には当該破損している分割データをパリティデータを用いて修復する。その後、これら分割データを合成し、ディスクに書き込む。従って、転送データが欠損した場合に、ミラー化せずに修復できるため、データ転送時の信頼性を保ちつつデータ転送の大容量化を図ることができる。

【 0 0 1 1 】

請求項 2 記載の発明では、上位ホストに送信する所定のデータをディスクから読み出すデータ読み出し工程と、当該読み出したデータを所定の処理をするデータ処理工程と、当該処理したデータを上位ホストに送信するデータ送信工程とを備えている。

【 0 0 1 2 】

そして、データ処理工程は、データ読み出し工程で読み出したデータを複数に分割すると共にパリティデータを生成するデータ分割工程と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納工程と、各分割データとパリティデータとをキャッシュモジュールから取り出すと共にいずれかの分割データが破損した場合に当該分割データをパリティデータを用いて修復するデータ修復工程と、これら分割データを合成するデータ合成工程とを備える、という構成を採っている。

【 0 0 1 3 】

このため、請求項 2 記載の発明では、前述した請求項 1 記載の発明と同様に転送データが欠損した場合に、ミラー化せずに修復できるため、データ転送時の信頼性を保ちつつデータ転送の大容量化を図ることができる。

【 0 0 1 4 】

請求項 3 記載の発明では、上位ホストからの指令によりディスクに対して所定のデータを書き込み及び読み出しすると共に、当該所定のデータを演算処理するアレイ制御部を備え、アレイ制御部は、所定のデータを少なくとも 2 つに分割すると共に当該所定のデータのパリティデータを生成するデータ分割機能と、分割データ 1 つが破損している場合にパリティデータを用いて当該破損している分割

データを修復する共に分割データを合成するデータ合成機能とを備える、という構成を採っている。

【0015】

このため、請求項3記載の発明では、所定のデータに基づくパリティデータが生成されると共に、このパリティデータを含めて所定のデータが少なくとも2つに分割され、分割されたデータの1つが破損している場合、修復機能によりパリティデータを用いて破損しているデータが修復される。従って、データ転送の信頼性を保ちつつ転送データをミラー化せず転送できるため、転送データの大容量化を図ることができる。

【0016】

請求項4記載の発明では、上位ホストからの指令によりディスクに対して所定のデータを書き込み及び読み出しすると共に、当該データを演算処理するアレイ制御部を備えている。

【0017】

そして、アレイ制御部は、所定のデータを少なくとも2つに分割すると共に当該所定のデータに基づいてパリティデータを生成するデータ分割部と、各分割データとパリティデータとをそれぞれ一時的に記憶する複数のキャッシュモジュールと、キャッシュモジュールの1つが故障した場合に当該故障のキャッシュモジュールに格納されていた分割データを残りの分割データとパリティデータとを用いて修復すると共に分割データを合成するデータ合成部とを備える、という構成を採っている。

【0018】

このため、請求項4記載に発明では、データ分割部により、所定のデータは分割され分割データが生成されると共に、当該所定のデータに基づくパリティデータが生成される。これら分割データとパリティデータは、複数のキャッシュモジュールに格納される。その後、これら分割データとパリティデータに基づいてデータ合成部によりデータが合成される。

【0019】

このとき、キャッシュモジュールの1つが故障し、分割データの1つが破損し

ている場合でも、データ合成部により残りの分割データとパリティデータに基づいて、破損された分割データを修復し、データを合成することができるため、データ転送の信頼性を保つことができる。さらに、1つのキャッシュモジュールが故障している場合でも、残りのキャッシュモジュールを用いてデータの転送を行うことができるため、I/O性能の低下を防ぐことができる。

【0020】

請求項5記載の発明では、請求項2記載のディスクアレイ装置において、各キャッシュモジュールの容量をそれぞれ等しく設定する、という構成を採っている。

【0021】

このため、請求項5記載の発明では、前述した請求項4記載の発明と同等の機能を有するほか、各キャッシュモジュールの容量がそれぞれ等しいため、分割データやパリティデータの生成を、キャッシュモジュールの容量をあまり考慮せず行えると共に、一種類のキャッシュモジュールの使用で済む。従って、分割データ及びパリティデータの生成方法が簡略化されると共に、製造コストを抑えることができる。

【0022】

請求項6記載の発明では、請求項4又は5記載のディスクアレイ装置において、各分割データ及びパリティデータの容量をそれぞれ等しく設定する、という構成を採っている。

【0023】

このため、請求項6記載の発明では、前述した請求項4又は5記載の発明と同等の機能を有するほか、各分割データ及びパリティデータの容量がそれぞれ等しいため、分割データ及びパリティデータの生成方法が簡略化される。従って、処理速度の向上を図ることができる。

【0024】

請求項7記載の発明では、請求項4、5又は6記載のディスクアレイ装置において、分割データ数と前記パリティデータ数との合計数を、前記キャッシュモジュール数に等しく設定する、という構成を採っている。

【0025】

このため、請求項7記載の発明では、前述した請求項4，5又は6記載の発明と同等の機能を有するほか、各キャッシュモジュールのそれぞれに分割データ又はパリティデータが格納されることになり、キャッシュモジュールの無駄がなくなり、キャッシュ容量を有効に利用することができる。

【0026】

請求項8記載の発明では、請求項4，5，6又は7記載のディスクアレイ装置において、分割データ数をキャッシュモジュール数より1つ少ない個数に設定する、という構成を採っている。

【0027】

このため、請求項8記載の発明では、前述した請求項4，5，6又は7記載の発明と同等の機能を有するほか、キャッシュモジュールの1つにパリティデータを割り当て、残りのキャッシュモジュールに分割データを割り当てることができる。従って、キャッシュメモリの容量をより有効にデータ転送に利用することができる。

【0028】

請求項9記載の発明では、ディスクに書き込まれる所定のデータを上位ホストから受け取った後、当該データを複数に分割すると共にパリティデータを生成するデータ分割処理と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納処理と、各分割データとパリティデータとをキャッシュモジュールから取り出すと共にいずれかの分割データに破損がある場合に当該破損している分割データをパリティデータとを用いて修復するデータ修復処理と、これら分割データを合成するデータ合成すると共に当該データをディスクに書き込むデータ合成処理とを備える、という構成を採っている。

【0029】

このため、請求項9記載の発明では、前述した請求項1，2又は3記載の発明と同様に転送データが欠損した場合に、ミラー化せずに修復できるため、データ転送時の信頼性を保ちつつデータ転送の大容量化を図ることができる。

【0030】

【発明の実施の形態】

以下、本発明の一実施形態を図 1 乃至図 2 に基づいて説明する。

【0031】

図 1 は本発明のディスクアレイ装置 1 の構成を示すブロック図である。この図 1 において、ディスクアレイ装置 1 は、上位ホスト 2 からの指令により複数の単体ディスク装置 4 に対して所定のデータを書き込み及び読み出しすると共に、当該所定のデータを分割及び合成するデータ分割機能及びデータ合成機能を有するアレイ制御部 5 を備えている。

【0032】

そして、アレイ制御部 5 は、前述した所定のデータを分割して分割データを生成すると共に当該所定のデータに基づいてパリティデータを生成するデータ分割部 7 と、各分割データとパリティデータとをそれぞれ一時的に記憶する複数のキャッシュモジュール 9, 10, 11 と、これらのキャッシュモジュール 9, 10, 11 に記憶された各分割データとパリティデータとに基づいてデータを合成するデータ合成部 8 とを備え、前述した所定のデータ用のデータ分割機能とデータ合成機能が実行されるようになっている。

【0033】

以下、これを詳述する。

【0034】

ディスクアレイ装置 1 には、前述したようにアレイ制御部 5 が備えられていて、このアレイ制御部 5 は、複数の単体ディスク装置 4 により構成される論理ディスク装置 3 に接続されている。

【0035】

具体的には、アレイ制御部 5 は、複数のアレイデータバス 13 を介して当該各アレイデータバス 13 に対応して備えられている複数のディスクインタフェース制御回路 14 に接続されている。そして、この各ディスクインタフェース回路 14 は、当該各ディスクインタフェース回路 14 に接続されているディスクインタフェース 15 を介して前述した単体ディスク装置 4 に接続されている。

【0036】

これにより、前述した上位ホスト 2 からの指令により、アレイ制御部 5 と単体ディスク装置 4 との相互間のデータの入出力が行われるようになっている。

【 0 0 3 7 】

また、アレイ制御部 5 は、前述した上位ホスト 2 に接続されている。具体的には、アレイ制御部 5 は、ホストデータバス 2 2 b を介してホストインタフェース制御回路 1 7 に接続され、当該ホストインタフェース制御回路 1 7 は、ホストインタフェース 1 8 を介して上位ホスト 2 に接続されている。

【 0 0 3 8 】

さらに、上記ホストインタフェース制御回路 1 7 は、マイクロプロセッサ制御回路 1 9 と接続されていて、上位ホスト 2 からの指令及びデータの入出力を制御している。なお、このマイクロプロセッサ制御回路 1 9 は、内部制御バス 2 0 を介してアレイ制御部 5 に接続されている。

【 0 0 3 9 】

ここで、アレイ制御部 5 には、前述したように、3 つのキャッシュモジュール 9, 1 0, 1 1 が備えられている。これら 3 つのキャッシュモジュール 9, 1 0, 1 1 は、容量が全て等しく形成されている。但し、これらのキャッシュモジュール 9, 1 0, 1 1 は、全て等しい容量に限定されるものではなく、それぞれ異なる容量であってもよい。また、当該キャッシュモジュールの個数は 3 つに限定されるものではない。

【 0 0 4 0 】

さらに、アレイ制御部 5 には、これらキャッシュモジュール 9, 1 0, 1 1 内のデータを制御するキャッシュ制御回路 6 が備えられている。そして、このキャッシュ制御回路 6 には、前述したデータ分割部 7 とデータ合成部 8 とが備えられている。

【 0 0 4 1 】

データ分割部 7 は、まず、上位ホスト 2 からホストデータバス 1 6 を介して又は単体ディスク装置 4 からアレイデータバス 1 3 を介して転送されてくる所定のデータを、複数の分割データブロック D 1, D 2 に分割すると共に、所定のデータに基づいてパリティデータ P を生成する。その後、内部データバス 1 2 を介し

て、キャッシュモジュール9に分割データブロックD1を、キャッシュモジュール10に分割データブロックD2を、キャッシュモジュール11にパリティデータPをそれぞれ格納するようになっている。

【0042】

データ合成部8は、上位ホスト2にホストデータバス16を介して又は単体ディスク装置4にアレイデータバス13を介して転送するための所定のデータを、内部データバス12を介してそれぞれキャッシュモジュール9、キャッシュモジュール10及びキャッシュモジュール11に格納されている複数のデータブロックD1、D2とパリティデータPとに基づいて合成するようになっている。

【0043】

ここで、前述した各分割データブロックD1、D2とパリティデータPとは、同一の容量に設定されている。これにより、分割データ及びパリティデータの生成方法が簡略化され、アレイ制御部5における処理の迅速化を図ることができる。

【0044】

また、前述したパリティデータは、キャッシュモジュールの個数より1少ない個数に分割されることが望ましい。すなわち、分割データとパリティデータとの合計数がキャッシュモジュールの個数に等しく設定されるとよい。これにより、キャッシュモジュールの個数をNとすると、実装キャッシュ容量の $(N-1)/N$ まで転送データの格納に割り当てることができ、転送データの大容量化を図ることができる。但し、前述した分割データの個数は、常にキャッシュモジュールの個数に依存するものではない。

【0045】

次に図2に基づいて、本実施形態の動作及びデータ読み書き制御方法について説明する。すなわち、本実施形態の動作によって、データ読み書き制御方法が実行されるようになっている。図2は、本実施形態の動作を表すフローチャートである。

【0046】

図2に示すように、データ読み書き制御方法は、ディスクに書き込まれる所定

のデータを上位ホストから受け取るデータ受け取り工程（ステップS1）と、当該受け取ったデータを所定の処理をするデータ処理工程（ステップS2, S3, S4, S5, S6, S7）と、当該処理したデータをディスクに書き込むデータ書き込み工程（ステップS8）とを備えている。

【0047】

また、ディスクからデータを読み出し、上位ホストに送信する場合には、上位ホストに送信する所定のデータをディスクから読み出すデータ読み出し工程と、当該読み出したデータを所定の処理をするデータ処理工程と、当該処理したデータを上位ホストに送信するデータ送信工程とを備えている。

【0048】

そして、データ処理工程（ステップS2, S3, S4, S5, S6, S7）は、データ受け取り工程（ステップS1）で受け取ったデータを複数に分割すると共にパリティデータを生成するデータ分割工程（ステップS2）と、この分割データとパリティデータとをそれぞれ個別にキャッシュモジュールに格納するデータ格納工程（ステップS3）と、各分割データとパリティデータとをキャッシュモジュールから取り出す（ステップS4）と共にいずれかの分割データが破損している場合（ステップS5）に当該分割データをパリティデータを用いて修復するデータ修復工程（ステップS6）と、これら分割データを合成するデータ合成工程（ステップS7）とを備えている。

【0049】

以下、これを詳述すると、データ受け取り工程は、上位ホスト2の指令により、単体ディスク装置4に書き込まれるデータ（ライトデータ）は、上位ホスト2からホストインタフェース18とホストインタフェース制御回路17とホストデータバス16とを介してアレイ制御部5に送信され、当該アレイ制御部は前記ライトデータを受け取る（ステップS1）。

【0050】

また、データ読み出し工程は、上位ホスト2の指令により、単体ディスク装置4から読み出されるデータ（リードデータ）は、当該単体ディスク装置4からディスクインタフェース15とディスクインタフェース制御回路14とアレイデー

タバス13とを介してアレイ制御部5に至る。

【0051】

そして、データ分割工程では、アレイ制御部5に至ったデータ（ライトデータ又はリードデータ）は、キャッシュ制御回路6内のデータ分割部7によって、複数の分割データブロックD1、D2に分割された後、パリティデータPも同時に生成される（ステップS2）。

【0052】

そして、データ格納工程では、当該データ分割部6により、内部データバス12を介して分割データブロックD1はキャッシュモジュール9に、分割データブロックD2はキャッシュモジュール10に、パリティデータPはキャッシュモジュール11にそれぞれ格納される（ステップS3）。但し、この格納箇所はこれに限定されるものではない。

【0053】

その後、キャッシュモジュール9、キャッシュモジュール10及びキャッシュモジュール11にそれぞれ格納されている複数のデータブロックD1、D2及びパリティデータPは取り出され（ステップS4）、いずれかの分割データが破損していなければ（ステップS5）、内部データバス12を介してデータ合成部8によりデータ合成工程で合成される（ステップS7）。

【0054】

合成されたデータは、データ書き込み工程において、上位ホスト2からの指示に従い、単体ディスク装置4に対する書き込みであれば、アレイ制御部5からアレイデータバス13とディスクインタフェース制御回路14とディスクインタフェース15とを介して論理ディスク装置11を構成する単体ディスク装置4に転送され、当該データは各単体ディスク4に書き込まれる（ステップS8）。

【0055】

また、合成されたデータが単体ディスク装置4からの読み出しであれば、データ送信工程において、アレイ制御部5からホストデータバス16とホストインタフェース制御回路17とホストインタフェース18とを介して上位ホスト2に転送される。

【0056】

ここで、キャッシュモジュール9が故障した場合（ステップS5）におけるデータの取り扱い方法について説明する。

【0057】

前述したように、キャッシュモジュール9には分割データブロックD1が、キャッシュモジュール10には分割データブロックD2が、キャッシュモジュール11にはパリティデータPがそれぞれ格納されている。このため、キャッシュモジュール9が故障した場合には、当該キャッシュモジュール9に格納されている分割データブロックD1は破棄される。

【0058】

このとき、残りのキャッシュモジュール10、11に格納されている分割データブロックD2及びパリティデータPは、前述したデータ合成部8により当該各キャッシュモジュール10、11内から内部データバス12を介して取り出される。

【0059】

そして、データ修復工程において、この分割データブロックD1とパリティデータPとに基づいて、当該データ合成部8により、分割データブロックD1が生成される。すなわち、キャッシュモジュール9の故障により破損している分割データブロックD1が修復される（ステップS6）。

【0060】

その後、前述したように、これら分割データがデータ合成部8により合成され（ステップS7）、合成されたデータは、各単体ディスク4に書き込まれ（ステップS8）、あるいは、合成されたデータは、上位ホストに転送される。

【0061】

また、故障したキャッシュモジュール9は、保守員等により交換される。当該キャッシュモジュール9の交換後、再び前述のようにデータ転送が行われる。但し、当該キャッシュモジュール交換前に置いて、残りの正常なキャッシュモジュールを用いてデータ転送を行うことができる。

【0062】

このようにすることにより、キャッシュモジュールの1つが故障し、分割データの1つが破棄された場合でも、残りの分割データとパリティデータに基づいて、データを合成することができる。このため、ディスクアレイ装置のデータ転送時の信頼性を保つことができる。

【0063】

また、前述したように、3つのキャッシュモジュール9、10、11を用いた場合、実装キャッシュ容量の3分の2を転送データの格納に使用できる。従って、キャッシュモジュールの個数を増やした場合、当該キャッシュモジュールの個数をNとすると、実装キャッシュ容量の $(N-1)/N$ まで転送データの格納に割り当てることができるため、キャッシュメモリの有効利用を図ることができ、転送データの大容量化を図ることができる。

【0064】

さらに、保守員等により故障したキャッシュモジュール交換前においても、残りのキャッシュモジュールを用いてデータの転送を行うことができる。従って、I/O機能が損なわれないため、ディスクアレイ装置のI/O性能の低下を防ぐことができる。

【0065】

【発明の効果】

以上説明したように、本発明のディスクアレイ装置によると、上位ホストからの指令によりディスクに対して所定のデータを書き込み及び読み出しすると共に、当該所定のデータを演算処理するアレイ制御部を備え、アレイ制御部は、所定のデータを少なくとも2つに分割すると共に当該所定のデータのパリティデータを生成するデータ分割機能と、分割データ1つが破損している場合にパリティデータを用いて当該破損した分割データを修復する共に分割データを合成するデータ合成機能とを備えたので、分割データの1つが破損した場合でも、合成機能によりパリティデータに基づいてデータを修復することができ、転送データをミラー化する必要がなく上位ホストとディスクとのデータ転送時におけるデータの欠損を抑制することができるため、データ転送の信頼性を保ちつつ転送データの大容量化を図ることができる、という従来にない優れた効果を有する。

【0066】

また、アレイ制御部は、所定のデータを少なくとも2つに分割すると共に当該所定のデータに基づいてパリティデータを生成するデータ分割部と、各分割データとパリティデータとをそれぞれ一時的に記憶する複数のキャッシュモジュールと、キャッシュモジュールの1つが故障した場合に当該故障のキャッシュモジュールに格納されていた分割データを残りの分割データとパリティデータとを用いて修復すると共に分割データを合成するデータ合成部とを備えたので、キャッシュモジュールの1つが故障し、分割データの1つが破損した場合でも、データ合成部により、残りの分割データとパリティデータに基づいて、破損した分割データを修復し、データを合成することができるため、データ転送の信頼性を保ちつつ残りのキャッシュモジュールを用いてデータ転送を行えるため、I/O性能の低下を防ぐことができ、さらに、キャッシュモジュール個数をNとすると、実装キャッシュ容量の $(N-1)/N$ まで転送データを格納できるため、転送データの大容量化を図ることができる、という従来にない優れた効果を有する。

【0067】

また、各キャッシュモジュールの容量をそれぞれ等しく設定した場合、あるいは各分割データ及びパリティデータの容量をそれぞれ等しく設定した場合には、各キャッシュモジュール又は各分割データ及びパリティデータの容量がそれぞれ等しいため、当該各データを各容量に基づいて分割データ及びパリティデータを生成する必要がなく、当該生成方法が簡略化され、処理速度の向上を図ることができると共に、一種類のキャッシュモジュールを使用するため、製造コストを抑えることができる、をという優れた効果を有する。

【0068】

また、分割データ数とパリティデータ数との合計数をキャッシュモジュール数に等しく設定した場合には、キャッシュモジュールの容量をより有効に利用できるように、キャッシュモジュールの無駄がなくなり、データ転送の効率化を図ることができる、という優れた効果を有する。

【0069】

さらに、分割データ数をキャッシュモジュール数より1つ少ない個数に設定し

た場合には、キャッシュモジュールの1つにパリティデータを割り当て、残りのキャッシュモジュールに分割データを割り当てることができるため、キャッシュメモリの容量をより有効にデータ転送に利用することができ、転送データの大容量化を図ることができる、という優れた効果を有する。

【図面の簡単な説明】

【図1】

本発明の一実施形態を示すブロック図である。

【図2】

図1に開示したディスクアレイ装置の動作を示すフローチャートである。

【図3】

従来例のディスクアレイ装置を示すブロック図である。

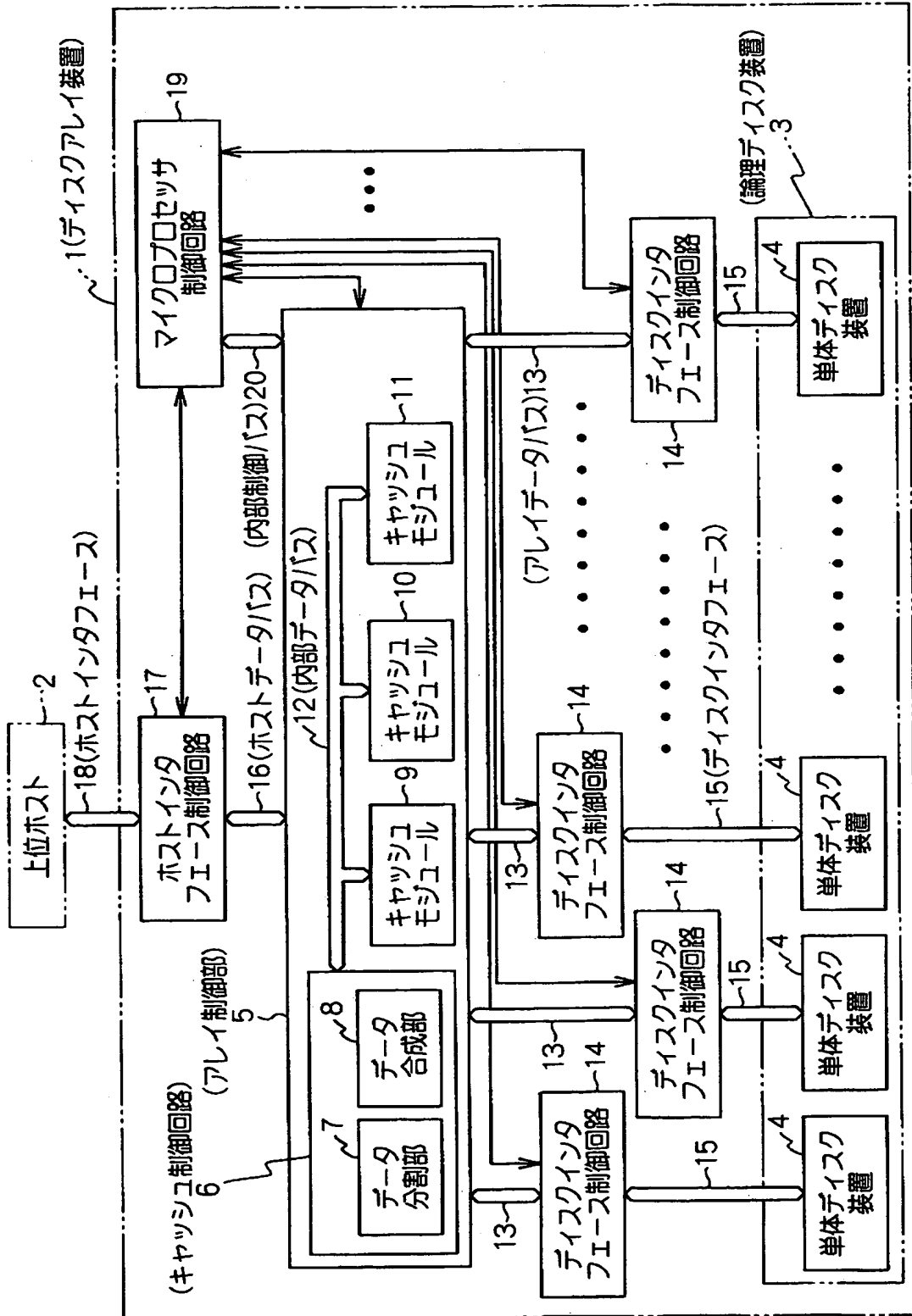
【符号の説明】

- 1 ディスクアレイ装置
- 2 上位ホスト
- 4 単体ディスク装置
- 5 アレイ制御部
- 7 データ分割部
- 8 データ合成部
- 9, 10, 11 キャッシュモジュール

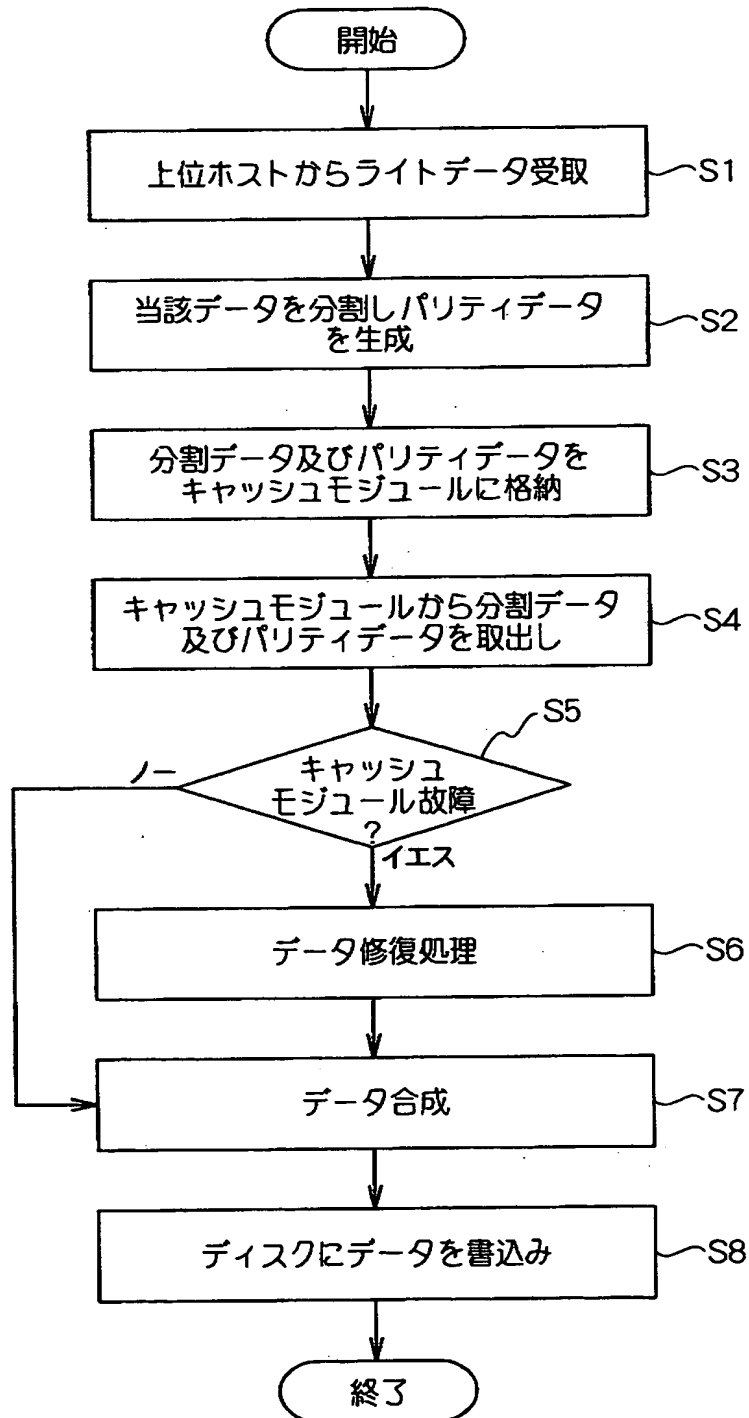
【書類名】

図面

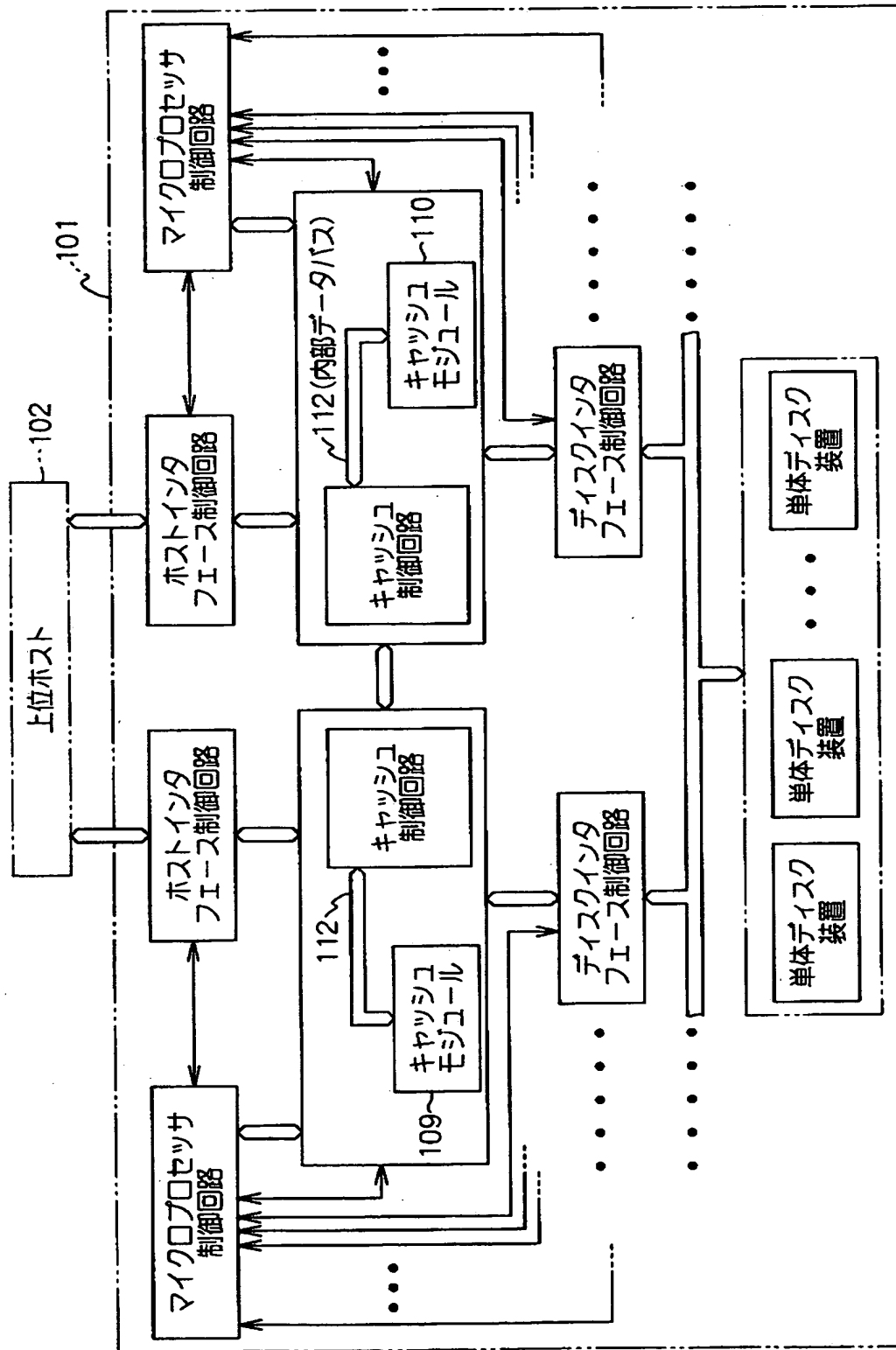
【図1】



【図 2】



【図3】



【書類名】 要約書

【要約】

【課題】 データ転送の信頼性を保つと共にデータ転送の大容量化を図るディスクアレイ装置を提供すること。

【解決手段】 ディスクに対して所定のデータを書き込み及び読み出しすると共に、当該データを演算処理するアレイ制御部 5 を備え、このアレイ制御部 5 は、所定のデータを少なくとも 2 つに分割すると共に当該所定のデータに基づいてパリティデータを生成するデータ分割部 7 と、各分割データとパリティデータとをそれぞれ一時的に記憶する複数のキャッシュモジュール 9, 10, 11 と、キャッシュモジュールの 1 つが故障した場合に当該故障のキャッシュモジュールに格納されていた分割データを残りの分割データとパリティデータとを用いて修復すると共に分割データを合成するデータ合成部 8 とを備えたこと。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日

[変更理由] 新規登録

住 所 東京都港区芝五丁目7番1号

氏 名 日本電気株式会社